Bad semidefinite programs: they all look the same

Gábor Pataki*

Abstract

We call a conic linear system P *badly behaved*, if for some objective function c the value $\sup \{ \langle c, x \rangle \mid x \in P \}$ is finite, but the dual program has no solution attaining the same value. We give simple, and exact characterizations of badly behaved conic linear systems. Surprisingly, it turns out that the badly behaved semidefinite system

$$x_1 \begin{pmatrix} \alpha & 1 \\ 1 & 0 \end{pmatrix} \preceq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{0.1}$$

where α is some real number, appears as a minor in all such systems in a well-defined sense. We prove analogous results for second order conic systems: here the minor to which all bad systems reduce is

$$x_1 \begin{pmatrix} \alpha \\ \alpha \\ 1 \end{pmatrix} \leq_K \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \tag{0.2}$$

where α is again some real number, and K is a second order cone.

Our characterizations provide $\mathcal{NP} \cap$co-$\mathcal{NP}$ certificates in the real number model of computing to recognize whether semidefinite, and second order conic systems are badly behaved.

Our main tool is one of our recent results that characterizes when the linear image of a closed convex cone is closed.

Key words: duality; closedness of the linear image; badly behaved instances; excluded minor characterizations

MSC 2010 subject classification: Primary: 90C46, 49N15; secondary: 52A40

OR/MS subject classification: Primary: convexity; secondary: programming-nonlinear-theory

1 Introduction. A sample of the main results

Conic linear programs, as introduced by Duffin ([17]) in the fifties provide a natural generalization of linear programming. Until the late eighties the main aspect of conic

*Department of Statistics and Operations Research, University of North Carolina at Chapel Hill

LPs that optimizers and convex analysts focused on was their duality theory. It is largely inherited from linear programming; however, when the underlying cone is not polyhedral, pathological phenomena occur, such as nonattainment of the optimal value, positive duality gaps, and infeasibility of the dual, when the primal program is bounded.

Nesterov and Nemirovskii [26] in the late eighties developed a general theory of interior point methods for conic LPs: the two most prominent efficiently solvable classes are semidefinite programs (SDPs), and second order conic programs (SOCPs). Today SDPs and SOCPs are covered in several surveys and textbooks, and are a subject of intensive research ([8, 5, 36, 1, 34, 14, 18]). The website [21] is a good resource to keep track of new developments.

This research was motivated by the curious similarity of pathological SDP instances appearing in the literature. We will present characterizations on when a conic linear system is well- or badly behaved from the standpoint of duality. Specialized to semidefinite programming, we obtain the main, surprising finding of the paper, namely that all badly behaved SDPs "look the same, " and the trivial badly behaved semidefinite system (0.1) appears as a minor in all such systems. We prove similar results for SOCPs.

We first present our general framework below.

Definition 1. Let X and Y be finite dimensional euclidean spaces, $c \in X$, $b \in Y$, and $A : X \to Y$ a linear map. Let $K \subseteq Y$ be a closed, convex cone, and write $s \leq_K t$ to denote $t - s \in K$. The conic linear programming problem parametrized by the objective function c is

$$\begin{array}{rl} \sup & \langle c, x \rangle \\ s.t. & Ax \leq_K b. \end{array} \tag{P_c}$$

The primal conic sytem is defined as

$$P = \{ x \mid Ax \leq_K b \}.$$

To define the dual program, we let A^* be the adjoint operator of A, and K^* the dual cone of K, i.e.,

$$K^* = \{ y \mid \langle y, x \rangle \geq 0 \, \forall x \in K \}.$$

The dual problem is then

$$\begin{array}{rl} \inf & \langle b, y \rangle \\ s.t. & y \geq_{K^*} 0 \\ & A^* y = c. \end{array} \tag{D_c}$$

Definition 2. We say that the system P is *well-behaved*, if for all c objective functions for which the value of (P_c) is finite, the dual has a feasible solution which attains the same value. We say that P is *badly behaved*, if it is not well-behaved. A more formal terminology, coined by Duffin, Jeroslow, and Karlovitz [16] for semi-infinite systems is saying that P yields, or does not yield *uniform LP-duality*; we use our naming for simplicity.

The most important class of conic LPs we study in this paper is semidefinite programs. We denote by $\mathcal{S}^n$ the set of n by n symmetric, and by $\mathcal{S}^n_+$ the set of of n by n symmetric positive semidefinite matrices. For $S, T \in \mathcal{S}^n$ let us write $S \preceq T$ to denote $T - S \in \mathcal{S}^n_+,$ and define their inner product as $S \bullet T = \sum_{i,j=1}^n s_{ij} t_{ij}.$ The primal-dual pair of semidefinite programming problems is then

$$
\begin{array}{lll}
 & \sup & \sum_{i=1}^m c_i x_i \\
(SDP_c) & s.t. & \sum_{i=1}^m x_i A_i \preceq B
\end{array}
\qquad\qquad
\begin{array}{lll}
\inf & B \bullet Y & \\
s.t. & Y \succeq 0 & (SDD_c) \\
 & A_i \bullet Y = c_i \ (i = 1, \dots, m), &
\end{array}
$$

where $A_1, \dots, A_m,$ and B are in $\mathcal{S}^n$. The pair $(SDP_c) - (SDD_c)$ is indeed a pair of conic LPs. So as not to distract from the main development of the paper, we formally show how they fit in the general framework of Definition 1 in the beginning of Section 3.

The primal semidefinite system is

$$P_{SD} = \{ x \mid \sum_{i=1}^m x_i A_i \preceq B \}.$$

The two motivating examples of pathological SDPs follow:

Example 1. In the problem

$$
\begin{array}{ll}
\sup & x_1 \\
s.t. & x_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \preceq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}
\end{array}
$$

the only feasible solution is $x_1 = 0$. In the dual program let us denote the components of Y by y_{ij}. The only constraint other than semidefiniteness is $2y_{12} = 1,$ so it is equivalent to

$$
\begin{array}{ll}
\inf & y_{11} \\
s.t. & \begin{pmatrix} y_{11} & 1/2 \\ 1/2 & y_{22} \end{pmatrix} \succeq 0,
\end{array}
$$

which has a 0 infimum, but does not attain it.

Example 2. The problem

$$
\begin{array}{ll}
\sup & x_2 \\
s.t. & x_1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \preceq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}
\end{array}
$$

again has an attained 0 supremum. The reader can easily check that the value of the dual program is 1, (and it is attained), so there is a finite, positive duality gap.

Curiously, while their "pathology" differs, Examples 1 and 2 still look similar: if we delete the second row and second column in all matrices in Example 2, remove the first matrix, and rename x_2 as x_1, we get back Example 1! This raises the following questions: Do all badly behaved semidefinite systems "look the same"? Does the system of Example 1 appear in all of them as a "minor"? In the following we make these questions more precise, and prove that the answer is yes to both. The paper contains a number of results for general conic LPs, but we state the ones for semidefinite systems first, since this can be done with minimal notation.

Assumption 1. *Theorems 1 and 2 are stated "modulo rotations", i.e., we assume that we can replace A_i by $T^T A_i T$ for $i = 1, \dots, m$, and B by $T^T B T$ for some invertible matrix T.*

We also assume that the maximum rank slack in P_{SD} is of the form

$$Z = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}. \tag{1.3}$$

Theorem 1. *The system P_{SD} is badly behaved, if and only if there exists a matrix V which is a linear combination of the A_i and B of the form*

$$V = \begin{pmatrix} V_{11} & e_1 & \dots \\ e_1^T & 0 & 0 \\ \vdots & 0 & V_{33} \end{pmatrix}, \tag{1.4}$$

where V_{11} is r by r, e_1 is the first unit vector, V_{33} is positive semidefinite (maybe an empty matrix), and the dots denote arbitrary components. □

The Z and V matrices provide a *certificate* of the bad behavior of P_{SD}. For example, they are

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, V = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

in Example 1, and

$$Z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, V = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

in Example 2. The reader may find it interesting to spot the certificates in the other badly behaved systems appearing in the literature.

Corollary 1. *Consider the following four elementary operations performed on P_{SD} :*

(1) Rotation: for some T invertible matrix, replace $A_1, \dots, A_m$ by $T^T A_1 T, \dots, T^T A_m T$, respectively, and B by $T^T B T$.

(2) Contraction: replace a matrix A_i by $\sum_{j=1}^m \lambda_j A_j$, and replace B by $B+\sum_{j=1}^m \mu_j A_j$, where the λ_j and μ_j are scalars, and $\lambda_i \neq 0$.

(3) Delete row ℓ and column ℓ from all matrices.

(4) Delete a matrix A_i.

If P_{SD} is badly behaved, then using these operations we can arrive at the system

$$x_1 \begin{pmatrix} \alpha & 1 \\ 1 & 0 \end{pmatrix} \preceq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

where α is some real number.

□

Well-behaved semidefinite systems have a similarly simple characterization:

Theorem 2. *The system P_{SD} is well-behaved, if and only if conditions (1) and (2) below hold:*

(1) there exists a matrix U of the form

$$U = \begin{pmatrix} 0 & 0 \\ 0 & I_{n-r} \end{pmatrix},$$

which satisfies

$$A_1 \bullet U = \ldots = A_m \bullet U = B \bullet U = 0.$$

(2) For all V matrices, which are a linear combination of the A_i and B and are of the form

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & 0 \end{pmatrix},$$

with $V_{11} \in \mathcal{S}^r$, we must have $V_{12} = 0$.

□

Example 3. The semidefinite system

$$x_1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \preceq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

is well-behaved, despite its similarity to the system in Example 1. The maximum slack and the U matrix of Theorem 2 are

$$Z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, U = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and since the entries in positions $(1,2)$ *and* $(1,3)$ *are zero in all constraint matrices, condition (2) trivially holds.*

Corollary 2. *The question*

"Is P_{SD} *well-behaved?"*

is in $\mathcal{NP} \cap \text{co-}\mathcal{NP}$ *in the real number model of computing.*

$\square$

Our main results build on three pillars. First, we describe characterizations (in conditions (2) and (3) in Theorem 5) of when the conic system P (defined in Definition 1) is well-behaved. These results involve the closedness of the linear image of K^* and $K^* \times \mathbb{R}_+$, where K^* is the dual cone of K. The closedness of these linear images is not easy to check in general, so we invoke a recent result from [31] on the closedness of the linear image of a closed convex cone, which we recap in Theorem 3. This accomplishes the following: with z a *maximum slack* in the system P, and $\mathrm{dir}(z, K)$ the set of feasible directions at z in the cone K, the well behaved nature of P will be related to how $\mathrm{dir}(z, K)$ intersects the linear subspace spanned by A and b. An equivalent characterization is given using the concept of strict complementarity, and the tangent space of K at z. Somewhat surprisingly, Theorem 5 unifies two seemingly unrelated conditions for P to be well-behaved: Slater's condition, and the polyhedrality of K.

Since $\mathrm{dir}(z, K)$ is a well structured set when K is the semidefinite or second order cone, a third step, which is some linear algebra, allows us to put the characterization for semidefinite and second order conic systems into a reasonably aesthetic shape, like Theorem 1 above. The closedness result from [31] also makes it possible to prove the polynomial time verifiability of the good and bad behavior of semidefinite and second order conic systems.

Literature review Linear programs with conic constraints were first proposed by Duffin in [17] mainly as a tool to better understand the duality theory of convex programs. For textbook and survey treatments on conic duality we refer to [5, 8, 35, 32, 25, 36, 14].

Our main results hinge on Theorem 3, which gives conditions for the linear image of a closed convex cone to be closed. While this question is fundamental, there is surprisingly little recent literature on the subject. We refer to [37] for a sufficient condition; to [4] for a necessary and sufficient condition on when the continuous image of a closed convex cone is closed, and to [3] for a characterization of when the linear image of a closed convex set is closed. More recently [7] studied a more general problem, whether the intersection of an infinite sequence of nested sets is nonempty, and gave a sufficient condition in terms of the sequence being *retractive*. The question whether a small perturbation of the linear operator still leads to the closedness of the linear image has been recently studied in [10].

The study of *constraint qualifications (CQs)* is also naturally related to our work. CQs are conditions under which a conic linear system, or a *semi-infinite linear system* (a system with finitely many variables, and infinitely many constraints), has desirable properties from the viewpoint of duality. Slater's condition (the existence of an x such that $b-Ax$ is in the relative interior of K) is sufficient for P to be well-behaved, but not necessary. The same holds for the polyhedrality of K. Another approach is to connect the desirable properties of a constraint set to the closedness of a certain related set. In this vein [16] introduced the notion of uniform LP-duality for a semi-infinite linear system, and showed it to be equivalent to such a CQ. Our condition (2) in Theorem 5 is analogous to Theorem 3.2 in [16]. More recently, [24, 13] and [23] studied such CQs, called *closed cone constraint qualifications* for optimization problems of the form $\inf \{ f(x) \,|\, x \in C, \, g(x) \in K \}$, where K is a closed convex cone, and C is a closed convex set. These CQs guarantee that the Lagrange dual of this problem has the same optimal value, and attains it. In particular, Condition (2) in our Theorem 5 specializes to Corollary 1 in [23] when K is the semidefinite cone, and Theorem 3.2 in [13] gives an equivalent condition, though stated in a different form. Closed cone CQs are weaker than Slater type ones, however, the closedness of the related set is usually not easy to check. A similar CQ for inequality systems in infinite dimensional spaces was given in [12], and a minimal cone based approach, which replaces K by a suitable face of K to obtain strong duality results was introduced in [11].

Where our results differ from the constraint qualifications listed above is that besides characterizing exactly when P is well-behaved, they are polynomial time verifiable for semidefinite and second order conic systems, and yield easily "viewable" certificates, as the Z and V matrices and the excluded system (0.1): the key is the use of Theorem 3.

Organization The rest of the introduction contains notation and definitions. In Section 2 we describe characterizations of when P is well behaved in Theorem 5. In Section 3 we prove Theorems 1 and 2, and Corollaries 1 and 2. In Section 4 we derive similar results for second order conic systems, and systems over p-cones. Finally, in Section 5 we prove that conic systems given in a different form (dual form, like the feasible set of (D_c)) or in the subspace form used by Nesterov and Nemirovskii in [26] are well-behaved, if and only if the the system we obtain by rewriting them in the standard form of P is.

Notation, and preliminaries The general references in convex analysis we used are: Rockafellar [33], Hiriart-Urruty and Lemaréchal [22], and Borwein and Lewis [9].

The closure, interior, boundary, relative interior and relative boundary of a convex set C are denoted by $\operatorname{cl} C$, $\operatorname{int} C, \operatorname{bd} C, \operatorname{ri} C$, and $\operatorname{rb} C$, respectively. For a convex set C, and $x \in C$ we define the following sets:

$$\begin{aligned} \operatorname{dir}(x, C) &= \{ y \,|\, x + \epsilon y \in C \text{ for some } \epsilon > 0 \}, && (1.5) \\ \operatorname{ldir}(x, C) &= \operatorname{dir}(x, C) \cap -\operatorname{dir}(x, C), && (1.6) \\ \tan(x, C) &= \operatorname{cl} \operatorname{dir}(x, C) \cap -\operatorname{cl} \operatorname{dir}(x, C). && (1.7) \end{aligned}$$

Here $\mathrm{dir}(x, C)$ is the set of feasible directions at x in C, $\mathrm{ldir}(x, C)$ is lineality space of $\mathrm{dir}(x, C)$, and $\tan(x, C)$ is the tangent space at x in C.

The range space, and nullspace of an operator A are denoted by $\mathcal{R}(A)$, and $\mathcal{N}(A)$, respectively. The linear span of set S is denoted by $\mathrm{lin}\, S$, and the orthogonal complement of $\mathrm{lin}\, S$ by $S^{\perp}$.

If x and y are elements of a euclidean space, we sometimes write x^*y for $\langle x, y \rangle$, and if $\lambda \in \mathbb{R}$, we write x/λ for $(1/\lambda)x$.

A set C is called a *cone,* if $\lambda x \in C$ holds for all $x \in C$, and $\lambda \geq 0$. If C is a convex cone, its dual cone is defined as

$$C^* = \{ y \mid \langle y, x \rangle \geq 0 \ \forall x \in C \}.$$

Suppose that C is a closed convex cone, and E is a convex subset of C. We say that E is a *face* of C, if $x_1, x_2 \in C$, and $1/2(x_1 + x_2) \in E$ implies that x_1 and x_2 are both in E. For an E face of C we have $E = C \cap \mathrm{lin}\, E$, and we define its *conjugate face* as

$$E^{\triangle} = C^* \cap E^{\perp}.$$

Similarly, the conjugate face of a G face of C^* is

$$G^{\triangle} = C \cap G^{\perp}.$$

Notice that the $()^{\triangle}$ operator is ambiguous, as it denotes two maps: one from the faces of C to the faces of C^*, and one in the other direction.

If C is a closed convex cone, and $x \in C$, we say that $u \in C^*$ is *strictly complementary to* x, if $x \in \mathrm{ri}\, E$ for some E face of C, (i.e., E is the smallest face of C that contains x, cf. Theorem 18.2 in [33]), and $u \in \mathrm{ri}\, E^{\triangle}$. Notice that u may be strictly complementary to x, but not the other way. The reason is that $(E^{\triangle})^{\triangle}$ is the smallest exposed face of C that contains E, i.e., the smallest face of C that arises as the intersection of C with a supporting hyperplane, and it only equals E, when E itself is exposed. In the semidefinite and second order cones all faces are exposed, so in these cones strict complementarity is a symmetric concept.

We say that a closed convex cone C is *nice*, if

$$C^* + E^{\perp} \text{ is closed for all } E \text{ faces of } C.$$

While this condition appears technical, it suffices to note that most cones appearing in the optimization literature, such as polyhedral, semidefinite, and second order cones are nice. For details, we refer to [31].

We state the following lemma for convenience:

Lemma 1. *Let C be a closed convex cone, $x \in C$, and E the smallest face of C that contains x. Then the following hold:*

$$\begin{aligned} \operatorname{dir}(x, C) &= C + \operatorname{lin} E, && (1.8)\\ \operatorname{ldir}(x, C) &= \operatorname{lin} E, && (1.9)\\ \operatorname{cl} \operatorname{dir}(x, C) &= (E^{\triangle})^*, && (1.10)\\ \tan(x, C) &= (E^{\triangle})^{\perp}. && (1.11) \end{aligned}$$

Proof Statements (1.8) and (1.10) are in Lemma 3.2.1 (Lemma 2.7 in the online version) in [29]. Statement (1.11) was also proved there, under the assumption that C is nice. In fact, it easily follows from (1.10) and (1.7) in general. We now prove (1.9). Let $y \in \operatorname{ldir}(x, C)$, then y and $-y$ are both in $\operatorname{dir}(x, C)$. Using (1.8), we have

$$y = c_1 + e_1 = -(c_2 + e_2)$$

for some $c_1, c_2 \in C$, e_1, $e_2 \in \operatorname{lin} E$. (Here e_i does not stand for a unit vector, rather for an element of E.) Hence

$$c_1 + c_2 = -(e_1 + e_2).$$

So $(1/2)(c_1 + c_2) \in C \cap \operatorname{lin} E = E$, and since E is a face, c_1 and c_2 are both in E, which completes the proof. $\square$

The following question is fundamental in convex analysis: under what conditions is the linear image of a closed convex cone closed? We state a slightly modified version of Theorem 1 from [31], which gives easily checkable conditions for the closedness of M^*C^*, which are "almost" necessary and sufficient.

Theorem 3. *Let M be a linear map, C a closed convex cone, and $x \in \operatorname{ri}(C \cap \mathcal{R}(M))$. Then conditions (1) and (2) below are equivalent to each other, and* necessary *for the closedness of M^*C^*.*

(1) $\mathcal{R}(M) \cap (\operatorname{cl} \operatorname{dir}(x, C) \setminus \operatorname{dir}(x, C)) = \emptyset$.

(2) There is $u \in \mathcal{N}(M^*) \cap C^*$ strictly complementary to x, and

$$\mathcal{R}(M) \cap (\tan(x, C) \setminus \operatorname{ldir}(x, C)) = \emptyset.$$

Furthermore, let E be the smallest face of C that contains x. If $C^* + E^{\perp}$ is closed, then conditions (1) and (2) are each sufficient *for the closedness of M^*C^*.*

$\square$

We make three remarks to explain Theorem 3 better:

Remark 1. For a set C let us write $\operatorname{fr}(C)$ for the *frontier of C,* i.e., the difference between the closure of C and C itself. Condition (1) in Theorem 3 relates the two questions:

$$\text{Is } \operatorname{fr}(M^*C^*) = \emptyset?$$

and

$$\text{Is } \mathcal{R}(M) \cap \operatorname{fr}(\operatorname{dir}(x, C)) = \emptyset?$$

The usefulness of Theorem 3 comes from the fact that $\operatorname{fr}(\operatorname{dir}(x, C))$ is a well-described set, when C is, and this well-described set helps us to understand the "messy" set M^*C^* better.

Remark 2. Theorem 3 is quite general: it directly implies the sufficiency of two classical conditions for the closedness of M^*C^*, and gives necessary and sufficient conditions for nice cones:

Corollary 3. *Let M and C be as in Theorem 3. Then the following hold:*

(1) If C is polyhedral, then M^*C^* is closed.

(2) If $\mathcal{R}(M) \cap \operatorname{ri} C \neq \emptyset$, then M^*C^* is closed.

(3) If C is nice, then conditions (1) and (2) in Theorem 3 are each necessary and sufficient for the closedness of M^*C^*.

Proof Let x and E be as in Theorem 3. If C is polyhedral, then so are the sets $C^* + E^\perp$, and $\operatorname{dir}(x, C)$, which are hence closed. So the sufficiency of condition (1) in Theorem 3 implies the closedness of M^*C^*.

If $\mathcal{R}(M) \cap \operatorname{ri} C \neq \emptyset$, then Theorem 6.5 in [33] implies $x \in \operatorname{ri} C$, hence $E = C$. Therefore $C^* + E^\perp = C^*$, and $\operatorname{dir}(x, C) = \operatorname{lin} C$, and both of these sets are closed. Again, the sufficiency of condition (1) in Theorem 3 implies the closedness of M^*C^*.

Finally, if C is nice, then $C^* + E^\perp$ is closed for *all* E faces of C, and this shows that conditions (1) and (2) in Theorem 3 are each necessary and sufficient for the closedness of M^*C^*. □

Remark 3. The second part of condition (2) in Theorem 3 is stated slightly differently in Theorem 1 in [31]: there it reads

$$\mathcal{R}(M) \cap ((E^\triangle)^\perp \setminus \operatorname{lin} E) = \emptyset.$$

However, $(E^\triangle)^\perp = \tan(x, C)$, and $\operatorname{lin} E = \operatorname{ldir}(x, C)$, as shown in Lemma 1.

In conic linear programs we frequently use the notion of solutions which are only "nearly" feasible. Here it suffices to define these only for (D_c). We say that $\{y_i\} \subseteq K^*$ is an *asymptotically feasible (AF)* solution to (D_c) if $A^*y_i \to c$, and the *asymptotic value of* (D_c) is

$$\operatorname{aval}(D_c) = \inf\{ \liminf_i b^*y_i \mid \{y_i\} \text{ is asymptotically feasible to } (D_c) \}.$$

Notice that if (D_c) is asymptotically feasible, then there is an AF solution $\{y_i\} \subseteq K^*$ with $\lim b^*y_i = \operatorname{aval}(D_c)$.

Theorem 4. *(Duffin [17]) Problem* (P_c) *is feasible with* $\mathrm{val}(P_c) < +\infty$, *iff* (D_c) *is asymptotically feasible with* $\mathrm{aval}(D_c) > -\infty$, *and if these equivalent statements hold, then*

$$\mathrm{val}(P_c) = \mathrm{aval}(D_c).$$

□

We denote vectors with lower case, and matrices and operators with upper case letters. The ith component of vector x is denoted by x_i, and the (i, j)th component of matrix Z by z_{ij}. We distinguish vectors and matrices of similar type with lower indices, i.e., writing $x_1, x_2, \ldots$ and $Z_1, Z_2, \ldots$. The jth component of vector x_i is denoted by $x_{i,j}$, and the (k, ℓ)th component of matrix Z_i by $z_{i,k\ell}$. This notation is somewhat ambiguous, as x_i may denote both a vector, or the ith component of the vector x, but the context will make it clear which one is meant. When a matrix Y is partitioned into blocks, these are denoted by Y_{ij}.

Adapting notation from [1], we use juxtaposing, or commas to denote concatenation of matrices along rows. That is, if $A_1, \ldots, A_k$ are matrices each with m rows, and $n_1, \ldots, n_k$ columns, respectively, then $(A_1 \ ... \ A_k) = (A_1, \ldots, A_k)$ has m rows, and $n_1 + \cdots + n_k$ columns. Also, we use semicolons to denote concatenation of matrices along columns. That is, if $A_1, \ldots, A_k$ are matrices each with n columns, and $m_1, \ldots, m_k$ rows, respectively, then $(A_1; \ldots; A_k) = (A_1^T, \ldots, A_k^T)^T$ is a matrix with n columns, and $m_1 + \cdots + m_k$ rows.

With some abuse of notation we similarly denote concatenation of elements of sets, and of operators. If $x_i \in X_i$ for sets X_i $(i = 1, \ldots, k)$, then $(x_1; \ldots; x_k)$ stands for the corresponding element of $X_1 \times \ldots \times X_k$. If X, Y, A and b are as in Definition 1, then the operator

$$\begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}$$

maps from $X \times \mathbb{R}$ to $Y \times \mathbb{R}$, by assigning $(Ax + bx_0; x_0)$ to $(x; x_0)$, and

$$\begin{pmatrix} A^* & 0 \\ b^* & 1 \end{pmatrix}$$

denotes its adjoint operator. If A_1 and A_2 are matrices, then we define

$$A_1 \oplus A_2 = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

and if X_1 and X_2 are sets of matrices, then we set

$$X_1 \oplus X_2 = \{ A_1 \oplus A_2 \,|\, A_1 \in X_1,\ A_2 \in X_2 \}.$$

For instance, $\mathcal{S}_+^r \oplus \{0\}$ (where the order of the 0 matrix will be clear from the context) stands for the set of matrices whose upper left r by r block is positive semidefinite, and the rest of the components are zero.

2 Characterizations of when P is well behaved

For the characterizations on the well behaved nature of $P,$ we assume that P is nonempty, and we need the following

Definition 3. *A* maximum slack *in P is a vector in*

$$\mathrm{ri}\,\{\, z \,|\, z = b - Ax,\ z \in K\,\} \;=\; \mathrm{ri}\,(\,(\mathcal{R}(A) + b) \cap K\,).$$

We will use the fact that for $z \in K$ the sets $\mathrm{dir}(z, K)$, $\mathrm{ldir}(z, K)$, and $\tan(z, K)$ (as defined in (1.5)) depend only the smallest face of K that contains z, cf. Lemma 1.

Theorem 5. *Let z be a maximum slack in P. Consider the statements*

(1) The system P is well-behaved.

(2) The set

$$\begin{pmatrix} A^* & 0 \\ b^* & 1 \end{pmatrix} \begin{pmatrix} K^* \\ \mathbb{R}_+ \end{pmatrix}$$

is closed.

(3) The set

$$\begin{pmatrix} A^* \\ b^* \end{pmatrix} K^*$$

is closed.

(4) $\mathcal{R}(A, b) \cap (\mathrm{cl}\,\mathrm{dir}(z, K) \setminus \mathrm{dir}(z, K)) = \emptyset.$

(5) There is $u \in \mathcal{N}((A, b)^*) \cap K^*$ strictly complementary to $z,$ and

$$\mathcal{R}(A, b) \cap (\tan(z, K) \setminus \mathrm{ldir}(z, K)) = \emptyset.$$

Among them the following relations hold:

$$\begin{array}{ccccc} (1) & \Leftrightarrow & (2) & \Leftarrow & (3) \\ & & \Downarrow & & \\ & & (4) & \Leftrightarrow & (5) \end{array}$$

In addition, let F be the smallest face of K that contains z. If the set $K^* + F^\perp$ is closed, then (1) through (5) are all equivalent.

□

We first remark that if P is a homogeneous system, i.e., $b = 0,$ then it is not difficult to see that P is well-behaved, if and only if A^*K^* is closed. So in this case Theorem 5 reduces to Theorem 3.

Second, we show an implication of Theorem 5: it unifies two classical, seemingly unrelated *sufficient* conditions for P to be well behaved, and gives *necessary and sufficient* conditions in the case when K is nice:

Corollary 4. *The following hold:*

(1) If K is polyhedral, then P is well behaved.

(2) If $z \in \operatorname{ri} K$, i.e., Slater's condition holds, then P is well behaved.

(3) If K is nice, then conditions (2) through (5) are all equivalent to each other, and with the well behaved nature of P.

Proof First, suppose that K is polyhedral. Then so is $K^* + F^\perp$, which is hence closed. The set $\operatorname{dir}(z, K)$ is also polyhedral, and hence closed. So the implication (4) $\Rightarrow$ (1) in Theorem 5 shows that P is well behaved. Next, assume $z \in \operatorname{ri} K$. Then F, the minimal face that contains z is equal to K. So $K^* + F^\perp$ equals K^*, and is hence closed, and $\operatorname{dir}(z, K) = \operatorname{lin} K$, which is also closed. Again, implication (4) $\Rightarrow$ (1) in Theorem 5 proves that P is well behaved. (In these two cases we do not use the opposite implication (4) $\Leftarrow$ (1).)

If K is nice, then $K^* + F^\perp$ is closed for *all* F faces of K, and this proves the last statement. $\square$

In preparation to prove Theorem 5, we define the sets

$$\begin{aligned} S_1 &= (\mathcal{R}(A) + b) \cap K, \\ S_2 &= \mathcal{R}(A, b) \cap K, \\ S_3 &= \mathcal{R}\begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} \cap \begin{pmatrix} K \\ \mathbb{R}_+ \end{pmatrix}. \end{aligned} \tag{2.12}$$

We need a somewhat technical lemma, whose proof is given in the Appendix.

Lemma 2. *Let $s_2 \in \operatorname{ri} S_2$, and $(s_3; s_0) \in \operatorname{ri} S_3$, where $s_3 \in K$, $s_0 \in \mathbb{R}_+$. Recall that z is the maximum slack in P, and F is the smallest face of K that contains z. Then we have*

(1) $s_2 \in \operatorname{ri} F$.

(2) $s_3 \in \operatorname{ri} F$, $s_0 > 0$.

(3) $\operatorname{dir}(z, K) = \operatorname{dir}(s_2, K) = \operatorname{dir}(s_3, K)$.

$\square$

Proof of Theorem 5: We use the notation S_1, S_2, and S_3 as introduced in (2.12). Also, in this proof, for a C closed convex cone, and $x \in C$ we will use the notation

$$\operatorname{frdir}(x, C) \;=\; \operatorname{cl} \operatorname{dir}(x, C) \setminus \operatorname{dir}(x, C). \tag{2.13}$$

The proof is somewhat redundant (i.e. we prove more implications than strictly necessary), but this should improve its readability.

Proof of (3) ⇒ (1) and (2) ⇒ (1) Let c be an objective vector, such that $c_0 := \operatorname{val}(P_c)$ is finite. Then $\operatorname{aval}(D_c) = c_0$ holds by Theorem 4, i.e., there is $\{y_i\} \subseteq K^*$ s.t. $A^* y_i \to c$, and $b^* y_i \to c_0$, in other words,

$$\begin{pmatrix} c \\ c_0 \end{pmatrix} \in \operatorname{cl} \begin{pmatrix} A^* \\ b^* \end{pmatrix} K^*. \tag{2.14}$$

By the closedness of the set in (2.14), there exists $y \in K^*$ such that $A^* y = c$, and $b^* y = c_0$. This argument shows the implication (3) ⇒ (1).

To prove (2) ⇒ (1), note that

$$\operatorname{cl} \begin{pmatrix} A^* \\ b^* \end{pmatrix} K^* \subseteq \operatorname{cl} \begin{pmatrix} A^* & 0 \\ b^* & 1 \end{pmatrix} \begin{pmatrix} K^* \\ \mathbb{R}_+ \end{pmatrix}. \tag{2.15}$$

By the closedness of the second set in (2.15) there exists $y \in K^*$, and $s \in \mathbb{R}_+$ such that $A^* y = c$, and $b^* y + s = c_0$. We must have $s = 0$, since $s > 0$ would contradict Theorem 4. Therefore y is a feasible solution to (D_c) with $b^* y = c_0$, and this completes the proof.

Proof of ¬(2) ⇒ ¬(1) Now, let c and c_0 be arbitrary, and suppose they satisfy

$$\begin{pmatrix} c \\ c_0 \end{pmatrix} \in \operatorname{cl} \begin{pmatrix} A^* & 0 \\ b^* & 1 \end{pmatrix} \begin{pmatrix} K^* \\ \mathbb{R}_+ \end{pmatrix}, \tag{2.16}$$

and

$$\begin{pmatrix} c \\ c_0 \end{pmatrix} \not\in \begin{pmatrix} A^* & 0 \\ b^* & 1 \end{pmatrix} \begin{pmatrix} K^* \\ \mathbb{R}_+ \end{pmatrix}. \tag{2.17}$$

Clearly, (2.16) is equivalent to the existence of $\{(y_i, s_i)\} \subseteq K^* \times \mathbb{R}_+$ s.t. $A^* y_i \to c$, and $b^* y_i + s_i \to c_0$. Hence $\operatorname{aval}(D_c) \leq c_0$. Also, Theorem 4 implies $\operatorname{val}(P_c) = \operatorname{aval}(D_c)$, so

$$\operatorname{val}(P_c) \;\leq\; c_0.$$

However, (2.17) implies that no feasible solution of (D_c) can have value $\leq c_0$. Hence either $\operatorname{val}(D_c) > c_0$ (this includes the case $\operatorname{val}(D_c) = +\infty$, i.e., when (D_c) is infeasible), or $\operatorname{val}(D_c)$ is not attained. In other words, c is a "bad" objective function.

Proof of $(2) \Rightarrow (4)$ **and of** $(2) \Leftrightarrow (4)$ **when** $K^* + F^\perp$ **is closed:** Let $(s_3; s_0) \in \operatorname{ri} S_3$, and G the smallest face of $K \times \mathbb{R}_+$ that contains $(s_3; s_0)$. Consider the statement

$$\operatorname{frdir}((s_3; s_0), K \times \mathbb{R}_+) \cap \mathcal{R}\begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} = \emptyset. \tag{2.18}$$

By Theorem 3 with

$$M = \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}, \; C = K \times \mathbb{R}_+, \; E = G$$

we have that (2) implies (2.18), and it is equivalent to it, when $(K \times \mathbb{R}_+)^* + G^\perp$ is closed. Obviously $(K \times \mathbb{R}_+)^* = K^* \times \mathbb{R}_+$. Next we claim

$$K^* \times \mathbb{R}_+ + G^\perp \text{ is closed} \;\Leftrightarrow\; K^* + F^\perp \text{ is closed}. \tag{2.19}$$

Indeed, by Lemma 2 we have $(s_3; s_0) \in \operatorname{ri}(F \times \mathbb{R}_+)$, so $G = F \times \mathbb{R}_+$. Using the fact that F and $\mathbb{R}_+$ are cones, we obtain $G^\perp = F^\perp \times \{0\}$, and this proves (2.19). (In fact, even if s_0 were 0, then $G^\perp = F^\perp \times \mathbb{R}$ would hold, and (2.19) would still be true, since all cones contained in the real line are closed.)

Next we claim

$$\operatorname{frdir}((s_3; s_0), K \times \mathbb{R}_+) = \operatorname{frdir}(s_3, K) \times \mathbb{R}. \tag{2.20}$$

Indeed, by definition, the set on the left hand side of (2.20) is the union of the two sets

$$\operatorname{frdir}(s_3, K) \times \operatorname{cl}\operatorname{dir}(s_0, \mathbb{R}_+)$$

and

$$\operatorname{cl}\operatorname{dir}(s_3, K) \times \operatorname{frdir}(s_0, \mathbb{R}_+).$$

But $s_0 > 0$ by Lemma 2, so $\operatorname{dir}(s_0, \mathbb{R}_+) = \operatorname{cl}\operatorname{dir}(s_0, \mathbb{R}_+) = \mathbb{R}$. So the second of these sets is empty, and the first is equal to the set on the right hand side of (2.20).

Given (2.20), statement (2.18) is equivalent to

$$\operatorname{frdir}(s_3, K) \cap \mathcal{R}(A, b) = \emptyset. \tag{2.21}$$

By (3) in Lemma 2 we obtain $\operatorname{frdir}(z, K) = \operatorname{frdir}(s_3, K)$, so (2.21) is equivalent to (4), and this completes the proof.

Proof of $(3) \Rightarrow (4) \Leftrightarrow (5)$ **and of** $(3) \Leftrightarrow (4) \Leftrightarrow (5)$ **when** $K^* + F^\perp$ **is closed:**

Let $s_2 \in \operatorname{ri} S_2$, and consider the statements (4-s_2) and (5-s_2) which are obtained from (4) and (5) by replacing z with s_2. By part (1) in Lemma 2 we have $s_2 \in \operatorname{ri} F$. Therefore Theorem 3 with $M = (A, b)$, $C = K$, $E = F$ proves

$$(3) \Rightarrow (4\text{-}s_2) \Leftrightarrow (5\text{-}s_2),$$

and that these statements are all equivalent, when $K^* + F^\perp$ is closed. By part (3) in Lemma 2, and the definition of the sets $\operatorname{ldir}(z, K)$ and $\tan(z, K)$ in (1.6)-(1.7) we have that (4-s_2) is equivalent to (4) and (5-s_2) to (5), and this completes the proof.

□

3 Characterizations of badly, and well-behaved semidefinite systems

We specialize the results on general conic systems in Theorem 5. The pair of SDPs $(SDP_c) - (SDD_c)$ fits into the general framework of Definition 1 by letting $X = \mathbb{R}^m$, $Y = \mathcal{S}^n$, defining the operator A as

$$Ax = \sum_{i=1}^{m} x_i A_i$$

for $x \in \mathbb{R}^m$, noting that

$$A^* Y = (A_1 \bullet Y, \ldots, A_m \bullet Y)^T$$

for $Y \in \mathcal{S}^n$, and that the set of positive semidefinite matrices is self-dual under the $\bullet$ inner product. Also, a maximum slack as defined in Definition 3 is a slack matrix with maximum *rank* in the semidefinite system P_{SD}, and the cone of positive semidefinite matrices is nice (see [31]).

To obtain Theorems 1 and 2 we use transformations of the A_i and B matrices. We need to make sure that these keep the maximum slack Z in the original form stated in (1.3). For ease of reference, we define what transformations are suitable below:

Definition 4. If $T = I_r \oplus P$, where P is an $n-r$ by $n-r$ invertible matrix, we call the transformation $T^T()T$ a *type 1 transformation.* If $T = Q \oplus I_{n-r}$, with Q an r by r orthonormal matrix, we call the the transformation $T^T()T$ a *type 2 transformation.*

We collect some results on the geometry of the semidefinite cone in Lemma 3. The proof is given in the Appendix.

Lemma 3. *Let Z be a positive semidefinite matrix of the form (1.3). Recall the definition of the set of feasible directions, and related sets from (1.5)-(1.7). Then*

$$\operatorname{ldir}(Z, \mathcal{S}^n_+) = \left\{ \begin{pmatrix} Y_{11} & 0 \\ 0 & 0 \end{pmatrix} \,\middle|\, Y_{11} \in \mathcal{S}^r \right\}, \tag{3.22}$$

$$\operatorname{cl}\operatorname{dir}(Z, \mathcal{S}^n_+) = \left\{ \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{12}^T & Y_{22} \end{pmatrix} \,\middle|\, Y_{22} \in \mathcal{S}^{n-r}_+ \right\}, \tag{3.23}$$

$$\tan(Z, \mathcal{S}^n_+) = \left\{ \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{12}^T & 0 \end{pmatrix} \,\middle|\, Y_{11} \in \mathcal{S}^r \right\}, \tag{3.24}$$

$$\operatorname{dir}(Z, \mathcal{S}^n_+) = \left\{ \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{12}^T & Y_{22} \end{pmatrix} \,\middle|\, Y_{22} \in \mathcal{S}^{n-r}_+, \mathcal{R}(Y_{12}^T) \subseteq \mathcal{R}(Y_{22}) \right\}. \tag{3.25}$$

Also, by using type 1 and type 2 transformations, all elements of $\operatorname{cl}\operatorname{dir}(Z, \mathcal{S}^n_+) \setminus \operatorname{dir}(Z, \mathcal{S}^n_+)$

can be brought into the form

$$V = \begin{pmatrix} V_{11} & e_1 & \dots \\ e_1^T & 0 & 0 \\ \vdots & 0 & V_{33} \end{pmatrix}, \tag{3.26}$$

where V_{11} is r by r, e_1 is the first unit vector, V_{33} is positive semidefinite (maybe an empty matrix), and the dots denote arbitrary components.

□

Given Lemma 3, the proof of Theorem 1 is straightforward.

Proof of Theorem 1 The equivalence (1) $\Leftrightarrow$ (4) in Theorem 5 shows that P_{SD} is badly behaved, iff there is a matrix V, which is a linear combination of the A_i and of B, and also satisfies

$$V \in \operatorname{cl} \operatorname{dir}(Z, \mathcal{S}^n_+) \setminus \operatorname{dir}(Z, \mathcal{S}^n_+).$$

By the last statement in Lemma 3 we have that all such V matrices can be brought to the form (3.26) with $V_{11} \in \mathcal{S}^r$, and V_{33} positive semidefinite (possibly an empty matrix). We need to apply these transformations to all the A_i matrices, and to B as well. Since they are all type 1, or type 2, they keep the maximum rank slack Z in the original form (1.3). □

Proof of Corollary 1: Suppose that P_{SD} is badly behaved. To find the minor system (0.1) we first add multiples of the A_i matrices to B to make B equal to the maximum slack, i.e.

$$B = Z = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}. \tag{3.27}$$

By Theorem 1 there is a matrix V of the form

$$V = \sum_{i=1}^{m} v_i A_i + v_0 B = \begin{pmatrix} V_{11} & e_1 & \dots \\ e_1^T & 0 & 0 \\ \vdots & 0 & V_{33} \end{pmatrix}, \tag{3.28}$$

with V_{11} being r by r, e_1 the first unit vector, V_{33} positive semidefinite (possibly an empty matrix), and the dots denoting arbitrary components. Since B is of the form described in (3.27), we may assume $v_0 = 0$, and we can also assume $v_1 \neq 0$ without loss of generality. We then replace A_1 by V, delete $A_2, \dots, A_m$, and all rows and columns except the first, and $(r+1)$st to arrive at the system

$$x_1 \begin{pmatrix} v_{11} & 1 \\ 1 & 0 \end{pmatrix} \preceq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

□

Proof of Theorem 2: We use the equivalence (1) $\Leftrightarrow$ (5) in Theorem 5. Let F be the smallest face of $\mathcal{S}^n_+$ that contains Z. Then clearly $F = \mathcal{S}^r_+ \oplus \{0\},$ and $F^{\triangle} = \{\,0\,\} \oplus \mathcal{S}^{n-r}_+.$

Hence a U positive semidefinite matrix is strictly complementary to Z if and only if

$$U \;=\; \begin{pmatrix} 0 & 0 \\ 0 & U_{22} \end{pmatrix}, \text{ with } U_{22} \succ 0.$$

So the first part of (5) in Theorem 5 translates into the existence of such a U which satisfies

$$A_1 \bullet U \;=\; \ldots \;=\; A_m \bullet U \;=\; B \bullet U \;=\; 0.$$

Let P be a matrix of suitably scaled eigenvectors of $U_{22},$ and $T = I_r \oplus P$. Replacing U by $T^T U T$, we may assume

$$U \;=\; \begin{pmatrix} 0 & 0 \\ 0 & I_{n-r} \end{pmatrix}.$$

Since for a W symmetric matrix we have $T^{-T} W T^{-1} \bullet T^T U T = W \bullet U$, we need to apply the inverse transformation $T^{-T}()T^{-1}$ to all A_i and B. Since $T^{-1} = I_r \oplus P^{-1},$ this is also a type 1 transformation, which leaves the maximum rank slack Z unchanged.

By (3.22) and (3.24) in Lemma 3 we have that the second part of condition (5) in Theorem 5 is equivalent to requiring that all V matrices, which are a linear combination of the A_i and B, and are of the form

$$V \;=\; \begin{pmatrix} V_{11} & V_{12} \\ V_{12}^T & 0 \end{pmatrix},$$

satisfy $V_{12} = 0$. $\square$

Proof of Corollary 2 The certificates of the *bad* behavior of P_{SD} are the Z and V matrices in Theorem 1. We can verify in polynomial time that Z is a maximum rank slack, using the algorithm of [11] or [30]. Since positive semidefiniteness of a matrix can be checked in polynomial time, we can also verify in polynomial time that a matrix V is of the form (1.4).

The well-behaved nature of P_{SD} can be confirmed via Theorem 2: checking that a matrix U satisfies condition (1) is trivial. Verifying Condition (2) amounts to showing the equality of two linear subspaces, which can be done by standard linear algebraic techniques. $\square$

4 A characterization of badly behaved second order conic systems

In this section we study badly behaved second order conic systems, and derive characterizations similar to Theorem 1 and Corollary 1. We first repeat the primal-dual pair of conic LPs here, for convenience:

$$
\begin{array}{llll}
 & \sup & c^*x \\
(P_c) & s.t. & Ax \leq_K b
\end{array}
\qquad\qquad
\begin{array}{llll}
\inf & b^*y & \\
s.t. & y \geq_{K^*} 0 & (D_c) \\
 & A^*y = c. &
\end{array}
$$

The second order cone in m-space is defined as

$$
\mathbb{SO}(m) \quad = \{\, x \in \mathbb{R}^m \,|\, x_1 \geq \sqrt{x_2^2 + \cdots + x_m^2} \,\}.
$$

The cone $\mathbb{SO}(m)$ is also self-dual with respect to the usual inner product in $\mathbb{R}^m$, and nice (see [31]). We call $(P_c) - (D_c)$ a primal-dual pair of second order conic programs (SOCPs), if

$$
K = K_1 \times \ldots \times K_t, \text{ and } K_i = \mathbb{SO}(m_i) \;\; (i = 1, \ldots, t),
$$

where m_i is a natural number for $i = 1, \ldots, t$. The primal second order conic system is

$$
P_{SO} \;=\; \{\, x \,|\, Ax \leq_K b \,\}.
$$

Here we view the operator A as simply a matrix with n columns. We first review two pathological SOCP instances. Despite the difference in their pathology (one's dual is unattained, but there is no duality gap, in the other there is a finite, positive duality gap, and the dual attains), the underlying systems look quite similar.

Example 4. Consider the SOCP

$$
\begin{array}{ll}
\sup & x_1 \\
s.t. & x_1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \leq_K \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.
\end{array}
\tag{4.29}
$$

with $K = \mathbb{SO}(3)$. Its optimal value is obviously zero. The dual problem is

$$
\begin{array}{ll}
\inf & y_1 + y_2 \\
s.t. & y_3 = 1 \\
 & y \geq_K 0.
\end{array}
\tag{4.30}
$$

The optimal value of the dual is $0,$ but it is not attained: to see this, first let

$$
y_i \;=\; (i + 1/i, -i, 1)^T \; (i = 1, 2, \ldots\,),
$$

then all y_i are feasible to (4.30), and their objective values converge to zero. However, $y = (y_1, -y_1, 1)$ is not in $K,$ no matter how y_1 is chosen, so there is no feasible solution with value zero.

Example 5. In this more involved example, which is a modification of Example 2.2 from [2], we show that the problem

$$\begin{array}{ll} \sup & x_2 \\ s.t. & x_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \leq_{K_1} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \\ & x_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \leq_{K_2} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \end{array} \tag{4.31}$$

where $K_1 = \mathbb{SO}(3)$, and $K_2 = \mathbb{SO}(2)$, and its dual have a finite, positive duality gap. The optimal value of (4.31) is clearly zero.

The dual is

$$\begin{array}{llll} \inf & y_1 + y_2 + w_1 & & \\ s.t. & y_1 + y_2 + w_1 - w_2 & = & 0 \\ & y_3 \qquad\qquad + w_2 & = & 1 \\ & y \in K_1, w \in K_2 & & \end{array} \tag{4.32}$$

Since y and w are in second order cones, $y_1 \geq -y_2$ and $w_1 \geq w_2$ must hold. Together with the first equation in (4.32) this implies

$$y_2 = -y_1, \tag{4.33}$$

$$w_2 = w_1. \tag{4.34}$$

Using (4.34), and the second constraint in (4.32)

$$y_3 = 1 - w_2 = 1 - w_1,$$

so we can rewrite problem (4.32) in terms of y_1 and w_1 only. It becomes

$$\begin{array}{llll} \inf & w_1 & & \\ s.t. & (y_1, -y_1, 1 - w_1)^T & \in & K_1 \\ & (w_1, w_1)^T & \in & K_2, \end{array}$$

and this form displays that the optimal value is 1.

Notice the similarity in the feasible sets of Examples 4 and 5: in particular, if we delete variable x_1 and the second set of constraints in the latter, we obtain the feasible set of Example 4!

We will use tranformations of the second order cone to make our characterizations nicer. It is well known that if C is a second order cone, then for all x_1 and x_2 in the interior of C there is a T symmetric matrix such that $Tx_1 = x_2$, and $TC = C$: such a rotation was used by Nesterov and Todd in developing their theory of self-scaled barriers for SOCP ([27, 28]). One can easily see that a similar T exists, if x_1 and x_2 are both nonzero, and on the boundary of C. So we are justified in making the following

Assumption 2. In P_{SO} we can replace A and b by TA and $Tb,$ respectively, where T is a block-diagonal matrix with symmetric blocks T_i and $T_i K_i = K_i$ for all $i = 1, \ldots, t.$ We also assume that the maximum slack in P_{SO} is of the form

$$z = \left(0; \ldots ; 0; \begin{pmatrix} 1 \\ e_1 \end{pmatrix} ; \ldots ; \begin{pmatrix} 1 \\ e_1 \end{pmatrix} ; e_1; \ldots ; e_1 \right), \tag{4.35}$$

(note that the dimension of the 0 and e_1 vectors can differ), and we let

$$\begin{array}{rcl} O & = & \{ i \in \{1, \ldots, t\} \,|\, z_i = 0 \}, \\ R & = & \{ i \in \{1, \ldots, t\} \,|\, z_i = (1; e_1) \}, \\ I & = & \{ i \in \{1, \ldots, t\} \,|\, z_i = e_1 \}. \end{array} \tag{4.36}$$

In the following we denote the blocks of a maximum slack z in P_{SO} by $z_1, \ldots, z_t,$ and write $v_1, \ldots, v_t$ for the corresponding blocks of a vector v with the same dimension as z. We also denote the columns of A by a_i $(i = 1, \ldots, n)$.

The main result on badly behaved second order conic systems follows.

Theorem 6. *The system P_{SO} is badly behaved, iff there is a vector $v = (v_1; \ldots ; v_t)$ which is a linear combination of the columns of A and of $b,$ and satisfies*

$$\begin{array}{rcl} v_i & \in & K_i \text{ for all } i \in O, \\ v_{i,1} & \geq & v_{i,2} \text{ for all } i \in R, \\ v_j & = & \begin{pmatrix} \alpha \\ \alpha \\ 1 \\ \vdots \end{pmatrix} \text{ for some } j \in R, \end{array} \tag{4.37}$$

where α is some real number, and the dots denote arbitrary components.

$\square$

The z slack, and the v vector in Theorem 6 give a certificate of the bad behavior of P_{SO}. They are

$$z = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \; v = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

in Example 4, and

$$z = \left(\begin{pmatrix} 1 \\ e_1 \end{pmatrix} ; \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right), \; v = \left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} ; \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$$

in Example 5.

Corollary 5. *Consider the following four elementary operations performed on* P_{SO} :

(1) Rotation: for some T *invertible matrix satisfying* $TK = K$, *replace* A *by* TA, *and* b *by* Tb.

(2) Contraction: replace a column a_i *of* A *with* $\sum_{j=1}^n \lambda_j a_j$, *and replace* b *by* $b + \sum_{j=1}^n \mu_j a_j$, *where the* λ_j *and* μ_j *are scalars, and* $\lambda_i \neq 0$.

(3) Delete rows other than a row corresponding to the first component of a second order cone from all matrices.

(4) Delete a column a_i.

If P_{SO} *is badly behaved, then using these operations we can arrive at the system*

$$x_1 \begin{pmatrix} \alpha \\ \alpha \\ 1 \end{pmatrix} \leq_K \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \tag{4.38}$$

where α *is again some real number, and* K *is a second order cone.*

$\square$

Corollary 6. *The question*

"Is P_{SO} *well-behaved?"*

is in $\mathcal{NP} \cap \text{co-}\mathcal{NP}$ *in the real number model of computing.*

$\square$

Proof of Theorem 6: Since the second order cone, and a direct product of second order cones is nice, we can specialize Theorem 5. By the equivalence (1) $\Leftrightarrow$ (4) we have that P_{SO} is badly behaved, iff there is a vector v, which is a linear combination of the columns of A and of b, and also satisfies

$$v \in \operatorname{cl} \operatorname{dir}(z, K) \setminus \operatorname{dir}(z, K). \tag{4.39}$$

Clearly, (4.39) holds, iff

$$\begin{array}{rcl} v_i & \in & \operatorname{cl} \operatorname{dir}(z_i, K_i) \text{ for all } i \in \{1, \dots, t\} \text{ and} \\ v_j & \notin & \operatorname{dir}(z_j, K_j) \text{ for some } j \in \{1, \dots, t\}. \end{array} \tag{4.40}$$

The form of z implies

$$\begin{array}{rcl} \operatorname{dir}(z_i, K_i) & = & K_i \text{ for all } i \in O, \\ \operatorname{dir}(z_i, K_i) & = & \mathbb{R}^{m_i} \text{ for all } i \in I, \end{array}$$

so in particular $\mathrm{dir}(z_i, K_i)$ is closed when $i \in O \cup I$. So (4.40) is equivalent to

$$
\begin{array}{rcl}
v_i & \in & K_i \text{ for all } i \in O, \\
v_i & \in & \mathrm{cl}\,\mathrm{dir}(z_i, K_i) \text{ for all } i \in R, \text{ and} \\
v_j & \in & \mathrm{cl}\,\mathrm{dir}(z_j, K_j) \setminus \mathrm{dir}(z_j, K_j) \text{ for some } j \in R.
\end{array} \tag{4.41}
$$

We now prove that if C is a second order cone, and $w = (1; e_1)$, then

$$
\mathrm{cl}\,\mathrm{dir}(w, C) = \{ v \mid v_1 \geq v_2 \}, \tag{4.42}
$$

$$
\mathrm{cl}\,\mathrm{dir}(w, C) \setminus \mathrm{dir}(w, C) = \{ v \mid v_1 = v_2,\ v_{3,:} \neq 0 \} \tag{4.43}
$$

holds, where $v_{3,:}$ denotes the subvector of v consisting of components $3, 4, \ldots$.

Let F be the minimal face of C that contains w, and $F^{\triangle}$ its conjugate face. Then it is easy to see that

$$
F^{\triangle} = \mathrm{cone}\,\{ (1; -e_1) \},
$$

and Lemma 1 implies

$$
\mathrm{cl}\,\mathrm{dir}(w, C) = (F^{\triangle})^*.
$$

Since $F^{\triangle}$ is generated by a single vector $(1; -e_1)$, its dual cone is the set of vectors having a nonnegative inner product with $(1; -e_1)$, i.e., (4.42) follows.

Next we prove (4.43). First note that (4.42) implies

$$
\mathrm{ri}\,\mathrm{dir}(w, C) = \mathrm{ri}\,\mathrm{cl}\,\mathrm{dir}(w, C) = \{ v \mid v_1 > v_2 \} \subseteq \mathrm{dir}(w, C). \tag{4.44}
$$

Suppose that v is in the left hand side set in equation (4.43). From (4.44) we get $v_1 = v_2$, and $v_{3,:} \neq 0$ follows directly from $v \notin \mathrm{dir}(w, C)$. Next suppose that v is in the set on the right hand side of equation (4.43). Then by (4.42) we have $v \in \mathrm{cl}\,\mathrm{dir}(w, C)$, and $v \notin \mathrm{dir}(w, C)$ follows directly from $v_{3,:} \neq 0$.

We then use (4.42) and (4.43), exchange components in v_j to make sure that $v_{j,3}$ is nonzero, and do a further rescaling to make sure that it equals 1. Modulo these transformation (4.41) is equivalent to (4.37). $\square$

The proof of Corollary 5 is analogous to the proof of Corollary 1, and the proof of Corollary 6 to the one of Corollary 2, so these are omitted.

The duality theory of conic programs stated over p-cones is by now well-understood [20, 2], and these optimization problems also find applications [15]. Hence it is of interest to characterize badly behaved p-conic systems, and this is the subject of the section's remainder. Letting p be a real number with $1 < p < +\infty$, the p-norm of $x \in \mathbb{R}^k$ is

$$
\| x \|_p = (|x_1|^p + \cdots + |x_k|^p)^{1/p},
$$

and the p-cone in dimension m is defined as

$$
K_{p,m} = \{ x \in \mathbb{R}^m \mid x_1 \geq \| (x_2, \ldots, x_m) \|_p \}.
$$

The dual cone of $K_{p,m}$ is $K_{q,m},$ where q satisfies $1/p + 1/q = 1.$

It is straightforward to see that $K_{p,m}$ does not contain a line, and is full-dimensional, and $w \in \operatorname{bd} K_{p,m} \setminus \{0\}$ iff

$$w \;=\; (\| x \|_p ; x)$$

for some nonzero $x \in \mathbb{R}^{m-1}$. For such a w we define its *conjugate vector* as

$$w^{\triangle} \;=\; (\| x \|_q ; -x).$$

Then $w^{\triangle} \in \operatorname{bd} K_{q,m} \setminus \{0\},$ and $\langle w, w^{\triangle} \rangle = 0$ hold.

Again recalling the conic programs $(P_c) - (D_c)$, we call them a pair of primal-dual p-conic programs, if

$$K = K_1 \times \ldots \times K_t, \text{ and } K_i = K_{m_i, p_i} \;\; (i = 1, \ldots, t),$$

where m_i is a natural number for $i = 1, \ldots, t,$ and $1 < p_i < +\infty$. The p-conic system is

$$P_p \;=\; \{ \, x \mid Ax \leq_K b \, \}.$$

When all p_i are $2,$ p-conic programs reduce to SOCPs.

Theorem 7. *Let z be a maximum slack in P_p, and define the index sets*

$$\begin{array}{rcl} O & = & \{ \, i \in \{1, \ldots, t\} \mid z_i = 0 \, \}, \\ R & = & \{ \, i \in \{1, \ldots, t\} \mid z_i \in \operatorname{bd} K_i \setminus \{0\} \, \}, \\ I & = & \{ \, i \in \{1, \ldots, t\} \mid z_i \in \operatorname{int} K_i \, \}. \end{array}$$

The system P_p is badly behaved, iff there is a vector $v = (v_1; \ldots; v_t)$ which is a linear combination of the columns of A and of $b,$ and satisfies

$$\begin{array}{rcl} v_i & \in & K_i \text{ for all } i \in O, \\ \langle v_i, z_i^{\triangle} \rangle & \geq & 0 \text{ for all } i \in R, \\ \langle v_j, z_j^{\triangle} \rangle \; = \; 0 \text{ and } v_j & \notin & \operatorname{lin}\{z_j\} \text{ for some } j \in R. \end{array}$$

$\square$

Theorem 7 reduces to Theorem 6, when all p_i are $2,$ and its proof is a straightforward generalization.

We finally remark that a characterization of well-behaved second-order and p-conic systems, analogous to Theorem 2 can be derived as well; also, the geometric and dual geometric cones of [19] have similar structure to p-cones, so Theorem 5 can be specialized to characterize badly and well-behaved systems over these cones.

5 Characterizing the well-behaved nature of conic systems in subspace, or dual form

In this section we look at conic systems stated in a form different from $P,$ and examine when they are well-behaved from the standpoint of duality. The short conclusion is that such a system is well-behaved, iff the system obtained by rewriting it in the "standard form" of P is. Most of the section's material is straightforward, so we only outline some of it.

We first define a primal-dual pair of conic linear programs in subspace form, parameterized by the primal objective y_0 as

$$\begin{array}{lllllll} & \inf & \langle y_0, z\rangle & \qquad & \inf & \langle z_0, y\rangle & \\ (P'_{y_0}) & s.t. & z \in (z_0 + L) \cap K & & s.t. & y \in (y_0 + L^{\perp}) \cap K^* & (D'_{y_0}), \end{array}$$

and the primal conic system in subspace form as

$$P' = (z_0 + L) \cap K.$$

Feasible solutions, boundedness, and the values of (P'_{y_0}) and (D'_{y_0}) are defined in the obvious way.

Definition 5. We say that the system P' is *well-behaved*, if for all y_0 objective functions for which the value of (P'_{y_0}) is finite, the problem (D'_{y_0}) has a feasible solution y that satisfies

$$\langle z_0, y\rangle + \mathrm{val}(P'_{y_0}) \;=\; \langle z_0, y_0\rangle.$$

We say that P' is *badly behaved*, if it is not well-behaved.

Let us recall the primal-dual pair $(P_c) - (D_c)$ and the primal conic system $P.$

Definition 6. *If*

$$L \;=\; \mathcal{R}(A) \text{ and } z_0 = b,$$

then we say that P and P' are equivalent. If in addition

$$A^* y_0 \;=\; c, \tag{5.45}$$

then we say that (P_c) and (P'_{y_0}), and (D_c) and (D'_{y_0}) are equivalent.

Clearly, for every conic system in the form of P there is an equivalent system in the form of $P',$ and vice versa. Also, if A is injective, which can be assumed without loss of generality, then for every conic LP in the form (P_c) there is an equivalent conic LP in the form of $(P'_{y_0}),$ and vice versa.

The following lemma is standard; we state, and prove it for the sake of completeness.

Lemma 4. *Suppose that* (P_c) *is equivalent with* (P'_{y_0})*, and* (D_c) *with* (D'_{y_0})*. If* x *is feasible to* (P_c) *with slack* z*, then* z *is feasible to* (P'_{y_0})*, and*

$$\langle c, x\rangle + \langle y_0, z\rangle \quad = \quad \langle y_0, z_0\rangle. \tag{5.46}$$

Also,

$$\mathrm{val}(P_c) + \mathrm{val}(P'_{y_0}) \quad = \quad \langle y_0, z_0\rangle. \tag{5.47}$$

Proof The first statement is obvious, and (5.46) follows from the chain of equalities

$$\begin{aligned}
\langle y_0, z_0\rangle &= \langle y_0, Ax + z\rangle \\
&= \langle y_0, Ax\rangle + \langle y_0, z\rangle \\
&= \langle A^* y_0, x\rangle + \langle y_0, z\rangle \\
&= \langle c, x\rangle + \langle y_0, z\rangle.
\end{aligned}$$

Suppose that x is feasible for (P_c) with slack z. Then z is feasible for (P'_{y_0}), and (5.46) holds, therefore

$$\langle c, x\rangle + \mathrm{val}(P'_{y_0}) \quad \leq \quad \langle y_0, z_0\rangle, \tag{5.48}$$

and since (5.48) is true for any x feasible solution to (P_c), the inequality $\leq$ follows in (5.47).

In reverse, let z be feasible to (P'_{y_0}). Then there is x that satisfies $z = z_0 - Ax$, so this x with the slack z is feasible to (P_c), so (5.46) holds, therefore

$$\mathrm{val}(P_c) + \langle y_0, z\rangle \quad \geq \quad \langle y_0, z_0\rangle. \tag{5.49}$$

Since (5.49) holds for any z feasible to (P'_{y_0}), the inequality $\geq$ follows in (5.47). $\square$

Combining Lemma 4 with Theorem 5 we can now characterize the well-behaved nature of P'.

Theorem 8. *Let* $z \in \mathrm{ri}\, P'$*. Consider the statements*

(1') The system P' *is well-behaved.*

(2') The set

$$\begin{pmatrix} I \\ -z_0^* \end{pmatrix} L^{\perp} + \begin{pmatrix} K^* \\ \mathbb{R}_+ \end{pmatrix}$$

is closed.

(3') The set

$$(L + z_0)^{\perp} + K^*$$

is closed.

(4') $\operatorname{lin}(L+z_0) \cap (\operatorname{cl}\operatorname{dir}(z,K) \setminus \operatorname{dir}(z,K)) = \emptyset$.

(5') There is $u \in (L+z_0)^\perp$ *strictly complementary to* z, *and*

$$\operatorname{lin}(L+z_0) \cap (\tan(z,K) \setminus \operatorname{ldir}(z,K)) = \emptyset.$$

Among them the following relations hold:

$$\begin{array}{ccccc} (1') & \Leftrightarrow & (2') & \Leftarrow & (3') \\ & & \Downarrow & & \\ & & (4') & \Leftrightarrow & (5') \end{array}$$

In addition, let F *be the smallest face of* K *that contains* z. *If the set* $K^* + F^\perp$ *is closed, then (1') through (5') are all equivalent.*

Proof Define the conic system P to be equivalent with P'. Then z is a maximum slack in P, and it suffices to show that statements (1), …(5) in Theorem 5 are equivalent to (1'), …(5') respectively.

Proof of $(1) \Leftrightarrow (1')$ Consider (P'_{y_0}) with some objective function y_0, and let $c = A^* y_0$. Then (P_c) and (P'_{y_0}) are equivalent, so (5.47) in Lemma 4 implies that $\operatorname{val}(P_c)$ is finite, if and only if $\operatorname{val}(P'_{y_0})$ is.

Also, (D_c) and (D'_{y_0}) are equivalent with the same feasible set, and the same objective value of a feasible solution. Combining this fact with (5.47), we see that there is a y feasible to (D_c) with $b^* y = \operatorname{val}(P_c)$, iff there is y is feasible to (D'_{y_0}) with $\langle z_0, y\rangle + \operatorname{val}(P'_{y_0}) = \langle y_0, z_0\rangle$, and this completes the proof.

Proof of $(2) \Leftrightarrow (2')$ Since $\mathcal{R}(A) = L$, we have $N(A^*) = L^\perp$. By a Theorem of Abrams (see e.g. [6]) if C is an arbitrary set, and M an arbitrary linear map, then MC is closed, if and only if $\mathcal{N}(M) + C$ is.

Applying this result we obtain

$$\begin{pmatrix} A^* & 0 \\ b^* & 1 \end{pmatrix} \begin{pmatrix} K^* \\ \mathbb{R}_+ \end{pmatrix} \ \textit{is closed} \Leftrightarrow \mathcal{N}\begin{pmatrix} A^* & 0 \\ b^* & 1 \end{pmatrix} + \begin{pmatrix} K^* \\ \mathbb{R}_+ \end{pmatrix} \ \textit{is closed.}$$

It is elementary to show that

$$\mathcal{N}\begin{pmatrix} A^* & 0 \\ b^* & 1 \end{pmatrix} = \begin{pmatrix} I \\ -z_0^* \end{pmatrix} L^\perp,$$

and this completes the proof.

The remaining equivalences follow from

$$\mathcal{R}(A,b) = \operatorname{lin}(L+z_0); \tag{5.50}$$

to prove (3) $\Leftrightarrow$ (3$'$) we also need to use Abram's theorem again. $\square$

In the rest of the section we discuss how to define the well-behaved nature of a conic system in dual form, and relate it to the well-behaved nature of the system obtained by rewriting it in subspace, or in primal form.

Definition 7. *We say that the system*

$$D = \{ y \in K^* \mid A^* y = c \}$$

is well-behaved, if for all b for which $v := \inf \{ b^* y \mid y \in D \}$ is finite, there is an x feasible solution to (P_c) with $c^* x = v$.

We can rewrite D in subspace form as

$$D' = (y_0 + L^{\perp}) \cap K^*,$$

where $L^{\perp} = \mathcal{N}(A^*)$, and $A^* y_0 = c$, and also in primal form as

$$D'' = \{ u \mid Bu \leq_{K^*} y_0 \},$$

where B is an operator with $\mathcal{R}(B) = \mathcal{N}(A^*)$. Actually, we have $D = D'$, but the well-behaved nature of D and D' are defined differently. Also note that $y \in D$ iff $y = y_0 - Bu$ for some $u \in D''$.

Theorem 9. *The systems D, D' and D'' are all well- or badly behaved at the same time.*

Proof Obviously, if $b = z_0$, then $\mathrm{val}(D'_{y_0}) = \inf \{ b^* y \mid y \in D \}$. By definition D' is well-behaved, if for all z_0 for which $\mathrm{val}(D'_{y_0})$ is finite, there is a z feasible solution to (P'_{y_0}) with

$$\langle z, y_0 \rangle + \mathrm{val}(D'_{y_0}) = \langle y_0, z_0 \rangle.$$

Using the first statement in Lemma 4 with (5.46) it follows that this holds exactly when there is an x feasible solution to (P_c) with $c^* x = \inf \{ b^* y \mid y \in D \}$. This argument proves that D is well-behaved if and only if D' is.

The argument used in the proof of the equivalence (1) $\Leftrightarrow$ (1$'$) in Theorem 8 shows that D' is well-behaved, if and only if D'' is, and this completes the proof.

$\square$

We can state characterizations of badly behaved semidefinite or second order conic systems given in a subspace, or dual form in terms of the original data. We give an example which is stated "modulo rotations", i.e. making Assumption 1. We omit the proof, since it is straightforward from Theorems 1 and 9.

Theorem 10. *Suppose that in the system*

$$Y \succeq 0,\ A_i \bullet Y \ =\ b_i\,(i = 1, \dots, m) \tag{5.51}$$

the maximum rank feasible matrix is of the form

$$\bar{Y} \ =\ \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Then (5.51) is badly behaved if and only if there is a matrix V and a real number λ with

$$A_i \bullet V \ =\ \lambda b_i \ (i = 1, \dots, m),$$

and

$$V \ =\ \begin{pmatrix} V_{11} & e_1 & \dots \\ e_1^T & 0 & 0 \\ \vdots & 0 & V_{33} \end{pmatrix},$$

where V_{11} is r by r, e_1 is the first unit vector, V_{33} is positive semidefinite (may be an empty matrix), and the dots denote arbitrary components. □

A Appendix

Proof of Lemma 2: We first prove

$$\operatorname{ri} S_1 \cap \operatorname{ri} S_2 \neq \emptyset. \tag{A.52}$$

Let $s_1 \in \operatorname{ri} S_1,\ s_2 \in \operatorname{ri} S_2,$ and let us write

$$s_1 \ =\ Ax_1 + b, \tag{A.53}$$
$$s_2 \ =\ Ax_2 + by_2 \tag{A.54}$$

for some x_1, x_2 and y_2. Since $s_1 \in S_2,$ by [33, Theorem 6.1] the line-segment $(s_1, s_2]$ is contained in $\operatorname{ri} S_2,$ in particular, for some small $\epsilon > 0$ we have

$$\begin{aligned} s_1(\epsilon) \ &:=\ (1-\epsilon)s_1 + \epsilon s_2 \\ &=\ A[(1-\epsilon)x_1 + \epsilon x_2] + b[(1-\epsilon) + \epsilon y_2] \ \in\ \operatorname{ri} S_2. \end{aligned}$$

Define

$$s_1'(\epsilon) \ =\ \frac{1}{1-\epsilon+\epsilon y_2} s_1(\epsilon). \tag{A.55}$$

If ϵ is sufficiently small, then the denominator in (A.55) is positive, so $s_1'(\epsilon) \in S_1$. Also, S_2 is a cone, and $s_1(\epsilon)$ is in $\operatorname{ri} S_2,$ hence so is $s_1'(\epsilon)$. Hence we have

$$s_1'(\epsilon) \in S_1 \cap \operatorname{ri} S_2. \tag{A.56}$$

Also, since $s_1'(\epsilon) \to s_1,$ as $\epsilon \to 0,$ and $s_1 \in \operatorname{ri} S_1,$ by (A.56) we obtain

$$s_1'(\epsilon) \in \operatorname{ri} S_1 \cap \operatorname{ri} S_2,$$

for small enough $\epsilon > 0,$ and this completes the proof of (A.52).

We note that since $\operatorname{ri} S_1$ is relatively open, and convex, by [33, Theorem 18.2] it is contained in the relative interior of a face of K. Since $\operatorname{ri} S_1$ and $\operatorname{ri} F$ intersect (in z), this face is F, so $\operatorname{ri} S_1 \subseteq \operatorname{ri} F$. Similarly, $\operatorname{ri} S_2$ is contained in the relative interior of a face of K; by (A.52) and $\operatorname{ri} S_1 \subseteq \operatorname{ri} F,$ we have $\operatorname{ri} S_2 \subseteq \operatorname{ri} F$. This finishes the proof of (1) in Lemma 2.

In (2) we first prove $s_0 > 0$. Since $(s_3; s_0) \in S_3,$ there exists x_3 satisfying

$$s_3 \quad = \quad Ax_3 + bs_0. \tag{A.57}$$

Let $s_1 \in S_1$. Expressing s_1 as in (A.53) shows $(s_1; 1) \in S_3$. Hence by [33, Theorem 6.4] the line-segment from $(s_1; 1)$ to $(s_3; s_0)$ can be extended past $(s_3; s_0)$ in S_3, i.e., for some small $\epsilon > 0$ we have

$$(1+\epsilon)(s_3; s_0) - \epsilon(s_1; 1) \quad \in \quad S_3, \tag{A.58}$$

hence $s_0 > 0$.

Define $s_3' := s_3/s_0$. Since S_3 is a cone, and $(s_3; s_0)$ is in $\operatorname{ri} S_3,$ so is $(s_3'; 1)$. So (A.58) holds with $(s_3'; 1)$ in place of $(s_3; s_0)$, i.e., there is a small $\epsilon > 0$ such that

$$(1+\epsilon)(s_3'; 1) - \epsilon(s_1; 1) \quad = \quad ((1+\epsilon)s_3' - \epsilon s_1; 1) \, \in \, S_3. \tag{A.59}$$

Dividing in (A.57) by s_0 we see that $s_3' \in S_1$. Also, (A.59) implies

$$(1+\epsilon)s_3' - \epsilon s_1 \, \in \, S_1,$$

hence $s_3' \in \operatorname{ri} S_1$. Since $\operatorname{ri} S_1 \subseteq \operatorname{ri} F,$ we obtain $s_3' \in \operatorname{ri} F$, hence $s_3 \in \operatorname{ri} F$, and this finishes the proof of (2).

Statement (3) then follows from using Lemma 1 with the fact that the minimal face of K that contains $z, s_2,$ and s_3 is the same, namely F. $\square$

Proof of Lemma 3: Let F be the smallest face of $\mathcal{S}_+^n$ that contains Z. Then clearly $F = \mathcal{S}_+^r \oplus \{0\},$ and $F^\triangle = \{\, 0 \,\} \oplus \mathcal{S}_+^{n-r}$. Hence $(F^\triangle)^*$ is the set on the right hand side in equation (3.23), and Lemma 1 with $C = \mathcal{S}_+^n$, $x = Z$, $E = F$ proves statements (3.22)-(3.24).

Next, fix $Y \in \operatorname{cl} \operatorname{dir}(Z, \mathcal{S}_+^n),$ and partition it as in the right hand side set in (3.23). Proving (3.25) is equivalent to verifying

$$Y \in \operatorname{dir}(Z, \mathcal{S}_+^n) \, \Leftrightarrow \, \mathcal{R}(Y_{12}^T) \subseteq \mathcal{R}(Y_{22}). \tag{A.60}$$

Let P be an orthogonal matrix whose columns are suitably scaled eigenvectors of Y_{22}, with the eigenvectors corresponding to zero eigenvalues coming first, and $T = I_r \oplus P$.

Clearly

$$T^TYT = \begin{pmatrix} Y_{11} & Y_{12}P \\ P^TY_{12}^T & P^TY_{22}P \end{pmatrix}.$$

We claim that both sides of the claimed equivalence (A.60) are unchanged, if we replace Y by T^TYT, i.e.,

$$\begin{aligned} Y \in \operatorname{dir}(Z, \mathcal{S}_+^n) \quad &\Leftrightarrow \quad T^TYT \in \operatorname{dir}(Z, \mathcal{S}_+^n), && \text{(A.61)} \\ \mathcal{R}(Y_{12}^T) \subseteq \mathcal{R}(Y_{22}) \quad &\Leftrightarrow \quad \mathcal{R}(P^TY_{12}^T) \subseteq \mathcal{R}(P^TY_{22}P) && \text{(A.62)} \end{aligned}$$

hold. Indeed, (A.61) follows from $T^TZT = Z$, and the definition of feasible directions. As to (A.62), the left hand side statement is equivalent to the existence of a matrix D such that

$$Y_{12}^T = Y_{22}D, \tag{A.63}$$

and the right hand side statement to the existence of a matrix D' such that

$$P^TY_{12}^T = P^TY_{22}PD'. \tag{A.64}$$

Now, if (A.63) is satisfied by some D, then (A.64) holds for $D' = P^{-1}D$. In reverse, if (A.64) holds for D', then $D = PD'$ verifies (A.63).

Let $s \geq 0$ be the number of nonzero eigenvalues of Y_{22}, and define $V = T^TYT$. Then V is of the form

$$V = \begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{12}^T & 0 & 0 \\ V_{13}^T & 0 & I_s \end{pmatrix}, \tag{A.65}$$

where $V_{11} = Y_{11}$, and $(V_{12}, V_{13}) = Y_{12}P$.

Clearly, $\mathcal{R}(P^TY_{12}^T) \subseteq \mathcal{R}(P^TY_{22}P)$ is equivalent to $V_{12} = 0$. Hence, using statements (A.61) and (A.62) the claimed equivalence (A.60) is equivalent to

$$V \in \operatorname{dir}(Z, \mathcal{S}_+^n) \Leftrightarrow V_{12} = 0. \tag{A.66}$$

Consider the matrix $Z + \epsilon V$ for some $\epsilon > 0$. If $V_{12} \neq 0$, then $Z + \epsilon V$ is not positive semidefinite for any $\epsilon > 0$, and this proves the direction $\Rightarrow$ in (A.66). As to $\Leftarrow$, if $V_{12} = 0$, then by the Schur-complement condition for positive semidefiniteness we have that $Z + \epsilon V \succeq 0$ iff

$$(I_r + \epsilon V_{11}) - (\epsilon V_{13})(\epsilon I_s)^{-1}(\epsilon V_{13}^T) \succeq 0,$$

and the latter is clearly true for some small $\epsilon > 0$. This argument finishes the proof of (A.66), and hence of (3.25).

We now prove the last statement of Lemma 3. The previous argument shows that all elements in $\operatorname{cl}\operatorname{dir}(Z, \mathcal{S}_+^n) \setminus \operatorname{dir}(Z, \mathcal{S}_+^n)$ can be brought to the form of (A.65) with $V_{11} \in \mathcal{S}^r$, $s \geq 0$, and $V_{12} \neq 0$ using a type 1 transformation.

Suppose that V is in this form. By exchanging rows and columns, we can make sure that the first column of V_{12} and the first row of V_{12}^T are nonzero, and by rescaling

that they have norm 1. These changes in V can clearly be done by suitable type 1 transformations. Next, let Q be an orthonormal matrix that maps the first column of V_{12} to e_1, and $T = Q \oplus I_{n-r}$. Replacing V by $T^T V T$ makes sure that it is in the form (3.26) with $V_{11} \in \mathcal{S}^r$, and V_{33} positive semidefinite (possibly an empty matrix). $\square$

Acknowledgement My sincere thanks are due to Shu Lu for her careful reading of the manuscript, and helpful comments. I also thank Minghui Liu for helpful comments.

References

[1] Farid Alizadeh and Donald Goldfarb. Second-order cone programming. *Math. Program. Ser. B*, 95:3–51, 2003.

[2] Erling D. Andersen, Cees Roos, and Tamás Terlaky. Notes on duality in second order and p-order cone optimization. *Optimization*, 51(4):627–643, 2002.

[3] Alfred Auslender. Closedness criteria for the image of a closed set by a linear operator. *Numer. Funct. Anal. Optim.*, 17:503–515, 1996.

[4] Heinz Bauschke and Jonathan M. Borwein. Conical open mapping theorems and regularity. In *Proceedings of the Centre for Mathematics and its Applications 36*, pages 1–10. Australian National University, 1999.

[5] Aharon Ben-Tal and Arkadii Nemirovski. *Lectures on modern convex optimization.* MPS/SIAM Series on Optimization. SIAM, Philadelphia, PA, 2001.

[6] Abraham Berman. *Cones, Matrices and Mathematical Programming.* Springer-Verlag, Berlin, New York, 1973.

[7] Dimitri Bertsekas and Paul Tseng. Set intersection theorems and existence of optimal solutions. *Math. Program.*, 110:287–314, 2007.

[8] Frédéric J. Bonnans and Alexander Shapiro. *Perturbation analysis of optimization problems.* Springer Series in Operations Research. Springer-Verlag, 2000.

[9] Jonathan M. Borwein and Adrian Lewis. *Convex Analysis and Nonlinear Optimization: Theory and Examples.* CMS Books in Mathematics. Springer, 2000.

[10] Jonathan M. Borwein and Warren B. Moors. Stability of closedness of convex cones under linear mappings. *J. Convex Anal.*, 16(3–4), 2009.

[11] Jonathan M. Borwein and Henry Wolkowicz. Regularizing the abstract convex program. *J. Math. Anal. App.*, 83:495–530, 1981.

[12] Jonathan M. Borwein and Henry Wolkowicz. A simple constraint qualification in infinite dimensional programming. *Math. Program.*, 35:83–96, 1986.

[13] Radu-Ioan Bot and Gert Wanka. An alternative formulation for a new closed cone constraint qualification. *Nonlin. Anal. Theor.*, 64:1367–1381, 2006.

[14] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[15] Samuel Burer and Jieqiu Chen. A p-cone sequential relaxation procedure for 0–1 integer programs. *Optim. Method. Softw.*, 24:523–548, 2010.

[16] Richard Duffin, Robert Jeroslow, and Les A. Karlovitz. Duality in semi-infinite linear programming. In *Semi-infinite programming and applications (Austin, Tex., 1981)*, volume 215 of *Lecture Notes in Econom. and Math. Systems*, pages 50–62. Springer, Berlin, 1983.

[17] Richard J. Duffin. Infinite programs. In A.W. Tucker, editor, *Linear Equalities and Related Systems*, pages 157–170. Princeton University Press, 1956.

[18] Osman Güler. *Foundations of Optimization.* Graduate Texts in Mathematics. Springer, 2010.

[19] Francois Glineur. Proving strong duality for geometric optimization using a conic formulation. *Ann. Oper. Res.*, 105(2):155–184, 2001.

[20] Francois Glineur and Tamás Terlaky. Conic formulation for l_p-norm optimization. *J. Optim. Theory Appl.*, 122(2):285–307, 2004.

[21] Christoph Helmberg. http://www-user.tu-chemnitz.de/˜semidef.html.

[22] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Convex Analysis and Minimization Algorithms.* Springer-Verlag, 1993.

[23] Vaitilingham Jeyakumar. A note on strong duality in convex semidefinite optimization: necessary and sufficient conditions. *Optim. Lett.*, 2(1):15–25, 2008.

[24] Vaitilingham Jeyakumar, N. Dinh, and G. M. Lee. A new closed cone constraint qualification for convex optimization. Technical Report AMR04/8, University of New South Wales, Sydney, Australia, 2004, 2004.

[25] Zhi-Quan Luo, Jos Sturm, and Shuzhong Zhang. Duality results for conic convex programming. Technical Report Report 9719/A, Erasmus University Rotterdam, Econometric Institute, The Netherlands, 1997.

[26] Yurii E. Nesterov and Arkady S. Nemirovski. *Interior Point Polynomial Algorithms in Convex Programming.* SIAM Publications. SIAM, Philadelphia, USA, 1994.

[27] Yurii E. Nesterov and Michael J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Math. Oper. Res.*, 22(1):1–42, 1997.

[28] Yurii E. Nesterov and Michael J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM J. Optim.*, 8:324–364, 1998.

[29] Gábor Pataki. The geometry of semidefinite programming. In Romesh Saigal, Lieven Vandenberghe, and Henry Wolkowicz, editors, *Handbook of semidefinite programming.* Kluwer Academic Publishers, also available from www.unc.edu/˜pataki, 2000.

[30] Gábor Pataki. A simple derivation of a facial reduction algorithm and extended dual systems. Technical report, Columbia University, 2000.

[31] Gábor Pataki. On the closedness of the linear image of a closed convex cone. *Math. Oper. Res.*, 32(2):395–412, 2007.

[32] James Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization*. MPS-SIAM Series on Optimization. SIAM, Philadelphia, USA, 2001.

[33] Tyrrel R. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, NJ, USA, 1970.

[34] Romesh Saigal, Lieven Vandenberghe, and Henry Wolkowicz, editors. *Handbook of semidefinite programming*. Kluwer Academic Publishers, 2000.

[35] Alexander Shapiro. Duality, optimality conditions, and perturbation analysis. In Romesh Saigal, Lieven Vandenberghe, and Henry Wolkowicz, editors, *Handbook of semidefinite programming*. Kluwer Academic Publishers, 2000.

[36] Michael J. Todd. Semidefinite programming. *Acta Numer.*, 10:515–560, 2001.

[37] Z. Waksman and M. Epelman. On point classification in convex sets. *Math. Scand.*, 38:83–96, 1976.